

## **National Informatics Standards**

### **NCI's Role as the Focus for the Needs of the Cancer Community**

For some years NCI has been building a compendium of cancer-related vocabulary and the technical resources to maintain and to disseminate the vocabulary. Standard vocabulary is an aspect of a larger set of standards that will be needed to provide cancer-related information and services in a structured form readily interpretable by people and computers.

**NCI must become a much more active voice in setting technical standards and promulgating cancer-related vocabulary.** The cancer community has requirements that must be addressed, and NCI is the natural focus for them.

More over, development of adequate standards is not enough. Adoption of standard techniques and vocabulary will provide benefit to the cancer community, but migration to them will cost money and will involve risks. **In certain critical cases, NCI should consider buying down the costs of migration and buffering the community from some associated risks.**

#### Background

In the emerging National Information Infrastructure, the unstructured information characteristic of the Web is giving way to structured information. Structured information is readily accessible to computers and people. The ability of computers to deal effectively with structured information will make it much easier for NCI and other health care and research-oriented organizations to use computers to retrieve, interpret and present actionable cancer-related information to colleagues and to the citizen.

Some of the core standards needed for structured content, such as eXtensible Markup Language (XML), are being built into the fabric of the Web. Other standards are being created to serve specific communities of Web users. In the health care and biomedical research community a number of standards development organizations (SDO) are defining how biomedical data should be structured, packaged for transmission, and documented. Other organizations are developing vocabulary to serve the health care and biomedical research community. Appendix 1 lists some of the more important standards development organizations and publishers of standard vocabularies. When mature, these standards and vocabularies will enable organizations that adhere to them to share information and to correctly interpret information from other community members.

**The needs of the cancer community must be addressed in these standards and vocabularies, but currently there is nothing to assure that the needs of the cancer community are being heard.**

NCI and the cancer community have begun to build large clinical and research collaborations and to establish business practices that depend upon information sharing. The standards and vocabulary we use to support these business practices and collaborations should conform to

available standards. That will make sharing information among collaborators and business partners easier to do using commercially developed products. However our collaborations and business practices require more than the standards and commercial vocabularies can deliver currently. Therefore our needs, and the solutions to them that we develop with our partners and collaborators, such as clinical trials vocabulary, should drive future development of the standards specific to the cancer community.

Consider the Mouse Models of Human Cancer Consortium as an example. The MMHCC is bringing together basic biologists, mouse researchers, genomics researchers and clinicians, specialists who do not share a common professional terminology. Successful collaboration among such a diverse group will require a common vocabulary. The Consortium will adopt vocabulary from existing sources where it is available and build it where it must. Once vocabulary is in hand, it must be made available to the membership, used in a consistent manner, and maintained as the Consortium moves ahead.

Vocabulary development and maintenance will represent a significant investment of the Consortium's resources. To the extent that the resulting vocabulary is incorporated into a standard vocabulary product, such as the ones listed in Appendix 1, the maintenance costs to the MMHCC will be reduced. Inclusion in a standard vocabulary product will also provide an avenue for the Consortium's terminology to enter common use.

The MMHCC is only one of a number of initiatives in which NCI is cooperating with external parties to produce mutually acceptable vocabulary. The Common Data Elements being developed by NCI in conjunction with CAPCURE, the breast and colorectal cooperative groups, by the Lung SPORE and CTEP is another. The ongoing development of the Common Toxicity Criteria and other vocabulary to meet the reporting needs of the cooperative groups and CTEP is yet another.

In fact, across the NCI initiatives are under way that will produce novel terminology. The Cancer Chromosome Aberration Project, Dr Buetow's informatics initiative, the Cancer Genome Anatomy Project, and cancer family registries, for example, will all doubtless generate vocabulary that will be of value to the cancer community.

NCI has begun to become a significant source of support for cancer-specific biomedical vocabulary and information standards driven collaboration. **NCI should therefore begin to formally and systematically participate in the standards and vocabulary development arena, serving as the focal point for the cancer community as a whole.**

#### What has been done to date

NCI's technology investments such as the Enterprise Vocabulary Server provide a means to store vocabulary, use it within NCI and make it available to a subset of the cancer community.

The Coding Subcommittee of the Extramural Information Systems Advisory Group (Barbara Spaholtz and Paulette Gray, chairs) is addressing the policies governing coding, the goals coding is to meet, and the management of vocabularies for coding in NCI systems.

NCI staff has contributed substantially to development of the cancer-related components of standard vocabularies. For example, SEER staff helped the World Health Organization to develop the ICD-O3 hematopoietic disease terminology, and CTEP staff assisted FDA to construct the oncology system organ class of what is now WHO's MedDRA vocabulary.

OI is discussing with the College of American Pathologists (CAP) establishment of a CRADA. The CRADA would build on the pending government-wide SNOMED/RT license that NLM is negotiating with CAP. The CRADA will provide a formal avenue to introduce cutting edge cancer-related vocabulary into SNOMED/RT.

NCI is represented, formally or informally on a number of SDOs and is partnering with other organizations. Examples include ongoing integration of the DCEs into the HCFA-sponsored metadata health standard, .X12N, and efforts to assure NCI is compliant with the HL7 version 3 messaging, data model and vocabulary standards.

However the wider cancer community's needs with regard to technical standards and vocabulary cannot be adequately addressed by internal NCI coordination, by collaborations between NCI and subsets of the cancer community, or by joint projects with individual software or vocabulary publishers.

#### What needs to be done

#### **1. NCI must become the focal point for assuring that the needs of the cancer community are heard in the national and international standards development process.**

- NCI should create a standing review panel to oversee cancer-specific information standards. The review panel should operate under one of the existing NCI advisory committees. The panel would peer review standards that are unique to the cancer community. The panel would advise NCI in regarding developing or promoting information standards for the cancer community. This panel should seek consensus from relevant stakeholders and liaison with federal agencies and organizations. Panel membership should be diverse, including practicing scientists and clinicians, organizations such as CAP/Cure, the Government-Computerized Patient Records Consortia (G-CPR), FDA and NLM, and HCFA, and prominent information researchers and industry representatives.
- NCI should ensure that our efforts to develop or promote cancer-specific information standards are coordinated with the broader health information standards. NCI should seek formal representation on SDOs including HL7, ISO TC215, ISO/IEC 11179 and ANSI HISB. NCI should strengthen its ties with WHO, CAP and other vocabulary publishers, as well as with federal organizations, such as HCFA and G-CPR that are engaged in standards development.
- When implementing standards, configuration management is vital to reduce risk and cost. NCI should provide the cancer community with services to manage and control change in cancer-specific standards and vocabulary. Vocabulary evolves over time, so means to continuously revise and control the quality of cancer-specific vocabulary must be provided to the community. Existing configuration management

procedures developed for the NCI Thesaurus, and the pending NCI/CAP CRADA, provide a point of departure. The cancer community will require an authoritative source for version control, seamless integration of updates and other services to successfully implement cancer specific standards. NCI should establish partnerships with other agencies, organizations and industry to meet this need.

2. **NCI should take the lead in involving the cancer community in the process of formal standards and vocabulary development.** The CDEs should be used as the initial context. Processes proven in the context of the CDEs should be used to address the community's needs for vocabulary and standards in other areas.
  - NCI should follow the recommendation of the Long Range Planning Committee (LRPC) by convening a national advisory meeting on oncology-related terminology and standards. The LRPC recommended Drs. Langholtz and Chute chair a working group to consider oncology-related terminology and standards. Membership would include outside experts and members of NCI staff. The working group would:
    - Identify oncology-related standards in which NCI ought to have a role, and oncology-specific standards that NCI should support throughout their lifecycle. (Oncology-related means a broader standard, such as HL7 version 3, which is relevant to oncology. Oncology-specific means a standard such as the CDEs or the common toxicity criteria, that was developed by and for the cancer community).
    - Institute change management processes for CDEs and other oncology-specific terminology. The existing NCI Thesaurus change management process and the pending NCI-CAP CRADA provide starting point. NCI should support development of guidelines and resource materials to assist the community to developing CDEs, and in designating "champions" for each round of CDE development.
    - Propose initiatives to augment the CDE information model. Meta-knowledge and meta-data about each concept is needed. Valid terms that denote the concept, interrelationships among concepts, and specification of the context in which a concept may be validly employed are needed. The working group will describe the CDE database schema and information model. Harmony between the CDE model and HCFA's meta-data repository will be a major goal. Existing NCI investments in description logic vocabulary development and the semantic net (part of the NCI EVS project) may be used to implement the model and schema.
    - Propose data dictionary entries for non-text data elements that will be in harmony with DICOM and other standards for representation of non-text information.
3. NCI has the major stake in the cancer community's adoption of standards-based collaboration and business practices. **NCI should take the lead in facilitating adoption by the community of cancer-specific standards and vocabulary.** NCI should address future concerns such as evolution of standards, and the financial and operational issues involved as the diverse cancer community adopts standard-based business practices and collaborations.
  - NCI should assure that proposed cancer-specific standards do not preclude smooth migration to new technologies. Generic information technology standards are frequently crafted by SDOs to avoid becoming a technological dead end.

Health-related standards are now also being crafted to avoid this problem. Cancer-specific standards that are conformant to such generic and health standards will benefit from the effort of the SDOs to provide future adaptability in the standards. Through participation in select generic SDOs and SDOs dealing with broad biomedical standards (some of which are enumerated in Appendix 1), NCI can help assure that the broad standards leave room for cancer-specific needs, and that cancer-specific standards possess extensibility and adaptability.

- Standards must be embodied in products for them to be beneficial. NCI should encourage a standards process in the cancer community that is like the one used by the Internet Engineering Task Force (IETF). In the IETF, not only must stakeholders be able to participate in definition of the proposed standards, but a working test bed and Internet-compatible infrastructure must be provided so that the standard can be widely tested. NCI should provide targeted support to test beds in areas that are critical to the cancer community.
- Organizations that choose to be early adopters of cancer-specific vocabulary and standards face costs and technical risks. NCI should target investment to assist early adopters to develop or purchase technologies embodying cancer-specific standards. Assistance could be through funding support or by making a standard available to the whole community at no cost or at reduced cost. The pending federal license of SNOMED/RT is an example.

4. **Diffusion of collaborative tools and new business process products to the general cancer community will require that NCI encourage awareness.** Success of standards-based products and services among the early adapters must be effectively communicated to the broader community, and impediments to other emulating the early adapters must be identified and dealt with.

- NCI should convene meetings and workshops at which the champions of standards, product vendors and early adapters interact with representatives of the broader cancer community.
- Generally new standards and products will be of interest to defined segments of the community, so these meetings would be focused on a solution to a problem common to that segment. The goal should be both to inform and to solicit feedback. These meetings should identify barriers to broad adoption and strategies to overcome them.

Appendix 1. National and International Standard Development Organizations Relevant to NCI

<b>Biomedical Standard</b>	<b>Development Organization</b>	<b>Purpose</b>
.X12N	ISO/IEC 11179, DOD/HCFA	Health Care Data Element Dictionary, metadata standard for biomedicine
In progress	ISO TC 215	Representation of health concepts within vocabularies
In progress	ANSI Health Information Standards Board	Healthcare records, data interchange, security, coding and terminology, representation of protocols, and more.
DICOM	National Electrical Manufacturers Association, Radiological Society of North America	Multimedia data representation, images in electronic medical record
In progress	Computerized Patient Records Institute, ASPE/DHHS	Patient privacy, confidentiality, HIPPA compliance
Health Level 7 version 3	ANSI accredited SDO	Reference Information Model, Code and Vocabulary API

<b>Vocabulary</b>	<b>Development Organization</b>	<b>Purpose</b>
MeSH	NLM	Indexing biomedical literature
SNOMED	College of American Pathologists, National Health Service	Pathology historically. Evolving into comprehensive medical vocabulary
ICD-10	WHO	In USA, HIPPA compliance
MedDRA	WHO	Regulatory reporting